Understanding hypervelocity sampling of biosignatures in space missions

Andres Jaramillo-Botero,[*,†] Morgan L. Cable,[‡] Amy E. Hofmann,[‡] Michael Malaska,[‡] Robert Hodyss,[‡] and Jonathan Lunine[¶]

†*Chemistry and Chemical Engineering Division, 1200 E California Blvd, California Institute of Technology*

‡*NASA Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109*

¶*Department of Astronomy and Carl Sagan Institute, Cornell University, 122 Sciences Drive, Ithaca, NY 14853*

E-mail: ajaramil@caltech.edu

Supporting Information

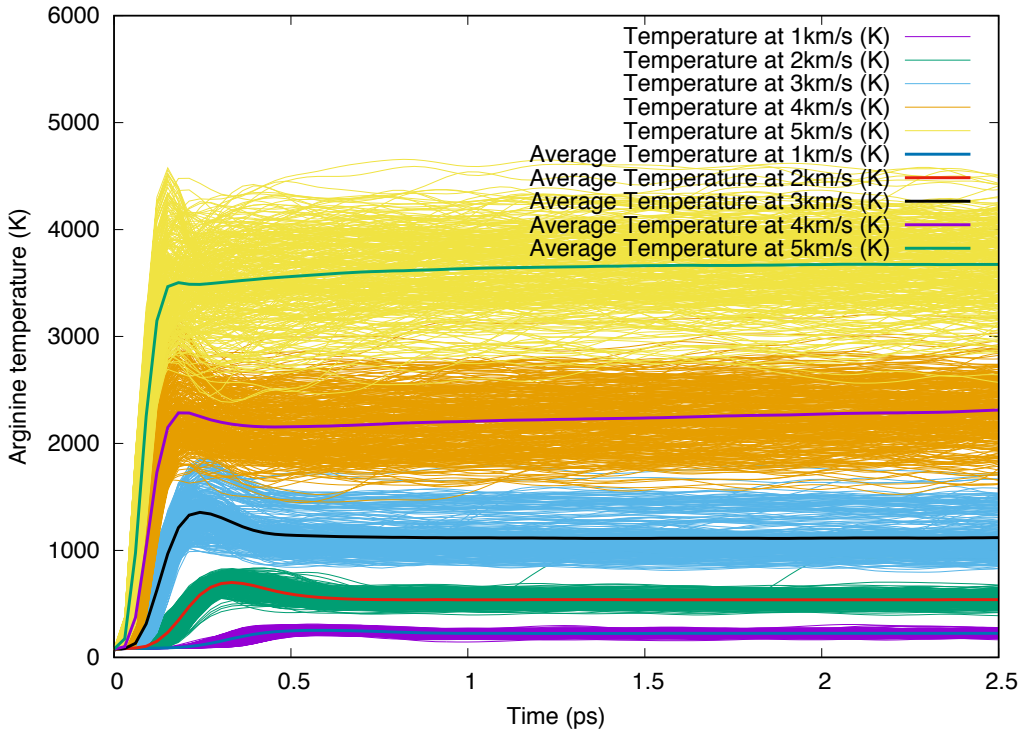


Figure S1: Arg thermalization on HVI ($v = 1 - 5km/s$ depicted) shows narrower standard deviation from average equilibrated temperatures, when compared to naphthalene due to its topology.

Figure S1, shows the HVI (up to 5 km/s) thermalization-induced case for the amino-acid Arg; a relatively (with respect to naphthalene) more flexible and elongated molecule, which shows a much narrower temperature distribution bandwidth per velocity due to its non-planar structure. $\Delta T_{3-4km/s}$ ≈900 K, albeit distributed among 26 atoms and 25 bonds (i.e. an average of 36K per bond, vs naphthalene, which gains an average 39.47K per bond).

Species of interest in this study are shown in Tables S1, S2, and S3.

Figure S4 the average bond energies/enthalpies at zero Kelvin for all the possible atomic pairwise bonds in the molecules involved in this study. bond energy is the average value for all the bond dissociation energies of all bonds of a certain type within a molecule, and it is strictly speaking, valid only for diatomic molecules.

Table S1: A selection of amino acids used in this study that could be an indicator for life if a few of them were to be quantified in Enceladus' plume (using 3-letter abbreviations).

Molecule	**Formula**	**Mass** (g/mol)	**Properties**
Ala	$C_3H_7NO_2$	89.09	Aliphatic, non-polar, hydrophobic, neutral
Arg	$C_6H_{14}N_4O_2$	174.2	Basic, basic polar, hydrophilic, charged (+)
Asp	$C_4H_7NO_4$	133.11	Acid, acidic polar, hydrophilic, charged (-)
Cit	$C_6H_{13}N_3O_3$	175.2	Polar, hydrophilic
Gly	$C_2H_5NO_2$	75.07	Aliphatic, non-polar, hydrophobic, neutral
Lys	$C_6H_{14}N_2O_2$	146.19	Basic, basic-polar, hydrophilic, charged (+))
Orn	$C_5H_{12}N_2O_2$	132.16	Aliphatic (not in proteins)
Ser	$C_3H_7NO_3$	105.09	Hydroxyl, polar, hydrophilic, neutral
Thr	$C_4H_9NO_3$	119.1192	Hydroxyl, polar, hydrophilic, neutral
Tyr	$C_9H_{11}NO_3$	181.19	Aromatic, polar, hydrophobic, neutral

Table S2: A selection of fatty acids used in this study that could be an indicator for life if a few of them were to be quantified in Enceladus' plume.

Molecule	**Formula**	**Mass** (g/mol)	**Un/Saturated**
Palmitic acid	$CH_3(CH_2)_{14}COOH$	256.430	Saturated
Cerotic acid	$CH_3(CH_2)_{24}COOH$	396.69	Saturated
Oleic acid	$CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$	282.468	Unsaturated

Table S3: Molecules associated with Titan' atmosphere in this study.

Name	**Formula**	**Mass** (g/mol)	**Structure/Functional group**
Napthalene	$C_{10}H_8$	128.1705	Aromatic
2-Octanone	$C_8H_{16}O$	128.215	Linear ketone
1,3-dicyanobenzene	$C_6H_4(CN)_2$	128.13	Aromatic nitrile

Table S4: Average bond energies in kcal/mol for the different types of bonds involved in this study, order in ascending sequence - adapted from.[1]

Bond	E_b	Ala	Arg	Asp	Cit	Gly	Lys	Orn	Ser	Thr	Tyr	Palmitic acid	Stearic acid	Arachidic acid	Oleic acid	Naphthalene	1,3-dicyanobenzene	2-octanone
C-N	71.7017	1	4	1	3	1	2	2	1	1	1							
C-C	83.6520	2	4	3	4	1	5	4	2	3	6	16	17	19	16	6	5	7
C-O	83.6520	1	1	1	1	1	1	1	2	2	1	1	1	1	1			
N-H	93.2122	2	6	2	6	2	4	4	2	2	2							
C-H	97.9924	4	7	3	7	2	9	7	3	5	7	33	35	39	33	8	4	16
O-H	109.9426	1	1	2	1	1	1	1	2	2	2	1	1	1	1			
C=C	146.0325										3				1	5	3	
C=N	146.9885		1															
C≡N	212.954		1														2	
C=O	178.0592	1	1	2	2	1	1	1	1	1	1	1	1	1	1			1
Total	bonds	12	26	14	24	9	23	20	13	16	23	52	55	61	53	19	14	24

Methods

QM calculations on non-adiabatic excitations in Figure 1 of the manuscript and for bond curves in Figure S9

The molecular geometry of CH_4 was optimized at M06-2X/6-311G**++ level (using Jaguar[2]), followed by a rigid scan of the C-H bond at the level of GVB-RCI with cc-pVTZ basis set (using GAMESS[3]).

The configuration interaction (CI) bond curves shown in Figure S9 were obtained using GAMESS63 with self-consistent field wavefunction (SCFTYP) at the level of restricted open shell Hartree-Fock with electronic correlations treated with the Ames Laboratory determinant full CI package (CITYP-ALDET) and the 6-G31 basis set for a rigid scan of the corresponding reaction coordinate at intervals of 0.05 Å.

Reactive molecular dynamics HVI protocol

The reaxFF reactive molecular dynamics simulation engine as implemented in the LAMMPS'[4] reax/c package,[5] along with a force field derived from the published work of Verlackt[6] and with C-C bond parameters from Liu[7] (without the dispersion correction) were used in this work.

Every target molecule was equilibrated to a temperature of 100 K, and then instantaneously accelerated to a constant velocity between, 1-12 km/s (with 1 km/s resolution), in the positive x direction (see Figure.**??**). A fixed wall was defined to bound the simulation domain on the $x+$ edge of the simulation box. This flat wall interacts with the atoms of the target molecule by generating a force on the atom in a direction perpendicular to the wall. The energy of wall-particle interactions was set to be a Lennard-Jones 12-6 potential of the form, $E = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$ where, r is the distance from the particle to the wall at position coordinate $x+$, ϵ is the strength factor for wall-particle interaction, σ is the size factor for wall-particle interaction, and r_c is the cutoff distance at which the particle and wall no longer interact, i.e. $r < r_c$ and the energy of the wall potential is shifted so that the wall-particle interaction energy is 0.0 past r_c. $\epsilon = 1.0$ kcal/mol, $\sigma = 1\mathring{A}$, and $r_c = 2.5\mathring{A}$ for all results reported here.

Each target molecule was minimized to an energy stopping tolerance of $1e-6$ or a force stopping tolerance of $1e-6$ kcal/mol-$\mathring{A}$, whichever occurred first. The molecule's center of mass was calculated and used as origin for a Cartesian coordinate system uvw attached to it; with all axis initially aligned with respect to the inertial xyz reference frame axes of the simulation. 180 sequential rotations about y followed by 180 sequential rotations about z were performed, at an interval resolution of 1 degree (2×180=360). After each orientation change, a temperature-controlled ramp was applied from 0-100 K in 1 ps, using a canonical NVT ensemble (fixed number of particles N, fixed volume V, and controlled temperature T) and an integration timestep of 1 fs. A Nosè-Hoover thermostat with a damping factor of 100 fs was used to control the temperature. The system was then thermally equilibrated at

100 K over 20 ps, verifying both final temperature and a normal distribution for all atomic velocity components. Once equilibrated, a restart file with atomic positions and velocities was written, for reuse at every subsequent impact angle (θ). The dynamics ensemble was changed to a microcanonical NVE (conserved energy) with an integration timestep of 0.1 fs for the impact simulations. An instantaneous change in molecular velocity along the positive x direction was applied, over the range $v = 1 - 12$ km/s, at 1 km/s intervals, over 30 ps of NVE dynamics using non-periodic and shrink-wrapped boundary conditions with a minimum value for all directions set to $\pm 20\mathring{A}$. This procedure was repeated for θ values between normal (0) and 85 degrees, at 5 degree intervals, with respect to the normal; in each case, starting from the restart file with the equilibrated system. To avoid memory allocation problems in the ReaxFF method, as implemented in the $reax/c$ pair style in LAMMPS, such as segmentation faults and bond check failed errors, that could occur for the highly compressed bond states expected during HVI impacts, we set a safezone of 2.0 and a mincap of 200. Non-bond neighbor lists were set to rebuild every 2 iterations.

For the ice-encapsulated cases, nanoclusters were equilibrated using an isothermal-isobaric NPT ensemble at 100 K and 1atm, for 5 ps with an integration timestep of 0.1 fs. A temperature damping factor of 100 time units and a pressure damping factor of 1000 time units were used for the Nose-Hoover thermostat and barostats. The equations of motion used are those of Shinoda et al,[8] which combine the hydrostatic equations of Martyna, Tobias and Klein[9] with the strain energy proposed by Parrinello and Rahman in.[10]

A thermodynamic data file was written out for each θ, ϕ, v case, at 100 step intervals, with every line containing step, temperature, total energy, potential energy, kinetic energy, and the individual components of the reaxFF energy expression, including charges computed from the charge equilibration method, every iteration with a lower cutoff of 0 and a higher cutoff of $10.0\mathring{A}$ and a precision tolerance set to 1x10e$-$6.

Computed center of mass positions, velocity and torque (Kcal/mole) vectors for every target organic molecule or fragments of it, were calculated and written to a file for every

θ, ϕ, v, every 100 iterations.

Trajectory files were dumped into flat files for every θ, ϕ, v case at 100 step intervals, as well as the chemical species information computed by the ReaxFF potential, as specified by *reax/c* pair style in LAMMPS. Averaged bond-order values per atom were computed using a bond-order cutoff of 0.3 to determine chemical bonds, every 100 timesteps, and chemical species information was written to a two-line output format filename. The first line containing species (chemical formulas) labels, and the second line consisting of timestep, total number of molecules, total number of distinct species, and number of molecules of each species.

Ab-initio predicted single electron impact Mass Spectra

Added energy from molecular ro-vibrational excitations during HVI for $v < v_f$ contribute to an effective 'thermalization' of the molecule, which unless dissipated through additional collisions, can affect (catalyze) the molecule's fragmentation pathways during the ionization stage of a mass spectrometer.

MS data is predicted from QM-based single-electron statistical models, adapted from,[11] involving the following steps:

1. Determining the average vibrational excitation energy gain after an impact and add it to the initial temperature of the vaporized substrate in the ionizing chamber of the MS,

2. Conformation search to produce a randomized ensemble of neutral ground state structures (M) using NVE dynamics and a thermostated temperature of 500 K,

3. Instantaneously ionize M (formally to the ground state of M*+), by assigning an ionization excess energy (IEE) to each random start structure, based upon a particular Poisson distribution, chosen so max IEE per atom is 0.6eV,

4. Propagate over the M+ PES and compute fragmentation PES pathways for each ionized structure on the fly, using semi-empirical AI-MD level theory,

5. Track secondary, tertiary, etc., fragmentations events systematically, and

6. Count fragments according to statistical weight, to produce calculated EI MS.

The computation of single electron impact MS requires knowledge of relative, energy-dependent chemical reaction rates (neglecting collisions, field effects, and photon exchange). After the primary electron impact an initial (e^-,$2e^-$) process leads to an excited ion state (M*+) that relaxes very quickly (0.5-2 ps) by internal conversion to a vibrationally hot ion ground state. Therefore, the nuclei can be propagated classically on a BornOppenheimer potential energy surface as determined via QM. Prior knowledge of the decomposition pathways is not needed. The simulation is guided by the gradient of the PES, generated on the fly by QM. The trajectories obtained by such calculations elucidate energetically accessible paths through a vast chemical reaction space of single-molecule decomposition reactions.

The semiempirical tight-binding QC DFTB+ with D3-dispersion correction was used as the main QM code `https://www.dftbplus.org` for these calculations, except for the IP/EA, which are needed to compute the charge distribution on fragments, and the MO calculations, which are needed to estimate the necessary ion (cation) state related quantities for the dynamic evolution of fragments. The IP/EA and MO calculations were performed using SCF PBE0/SV(P) in ORCA `https://orcaforum.kofo.mpg.de/app.php/portal`. A total of 450 (25×18) trajectories were used for the production run, with a maximum time of $2.5 - 5ps$.

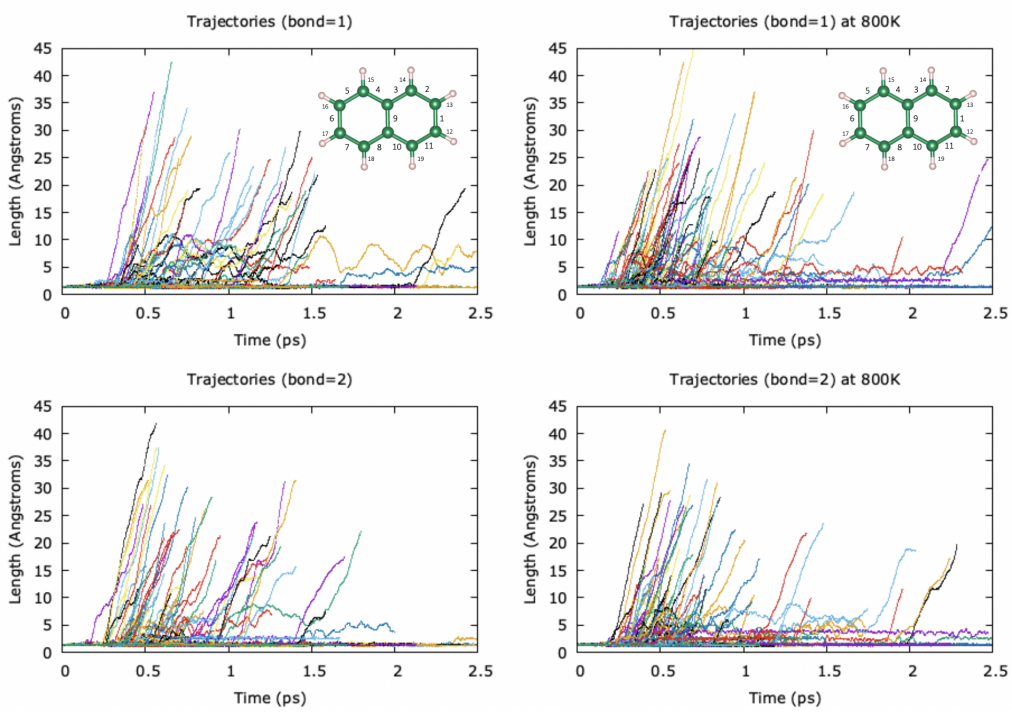


Figure S2: Bond trajectories (1-2) as a function of time after instantaneous ionization (up to 3 levels of cascaded fragmentation) of naphthalene at 500 K (left column) and 800 K (right column) vaporized substrate temperature. Bond length excursions beyond the equilibrium values (1.37-1.42$\mathring{A}$) indicate a dissociated bond.

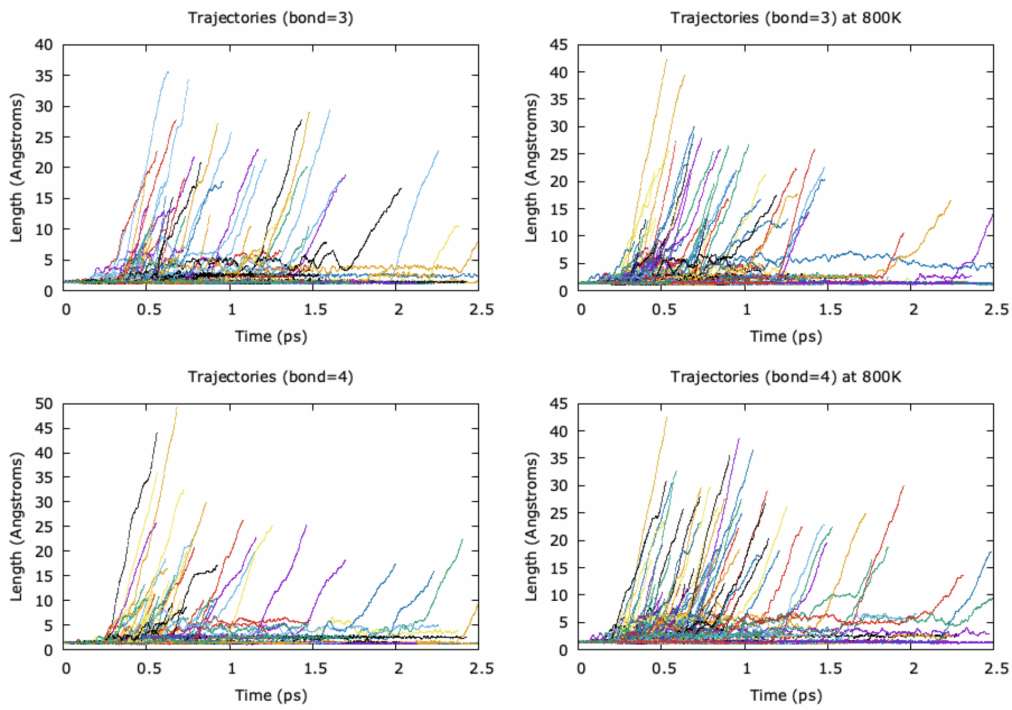


Figure S3: Bond trajectories (3-4) as a function of time after instantaneous ionization (up to 3 levels of cascaded fragmentation) of naphthalene at 500 K (left column) and 800 K (right column) vaporized substrate temperature. Bond length excursions beyond the equilibrium values (1.37-1.42$\mathring{A}$) indicate a dissociated bond.

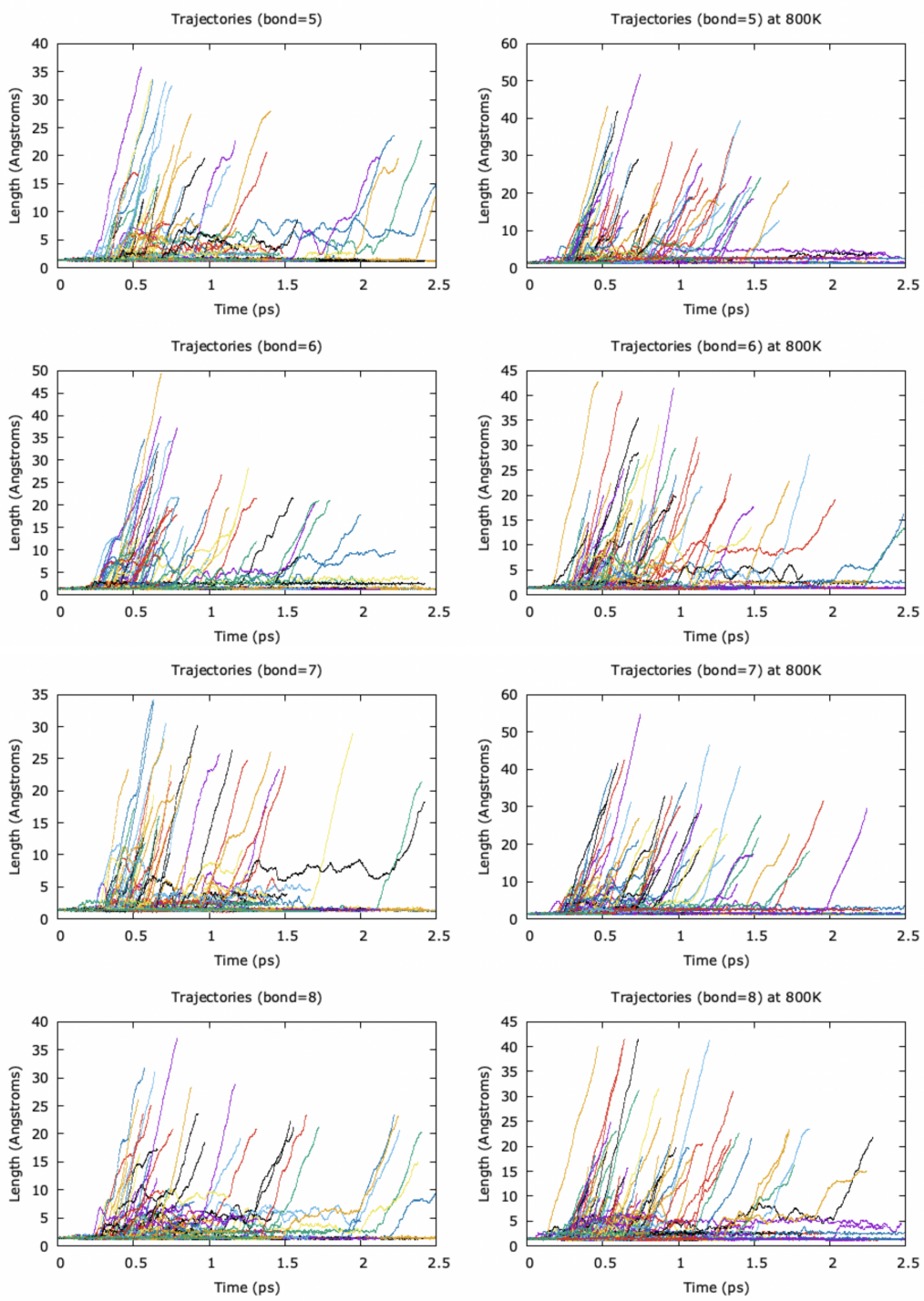


Figure S4: Bond trajectories (6-8) as a function of time after instantaneous ionization (up to 3 levels of cascaded fragmentation) of naphthalene at 500 K (left column) and 800 K (right column) vaporized substrate temperature. Bond length excursions beyond the equilibrium values (1.37-1.42$\mathring{A}$) indicate a dissociated bond.

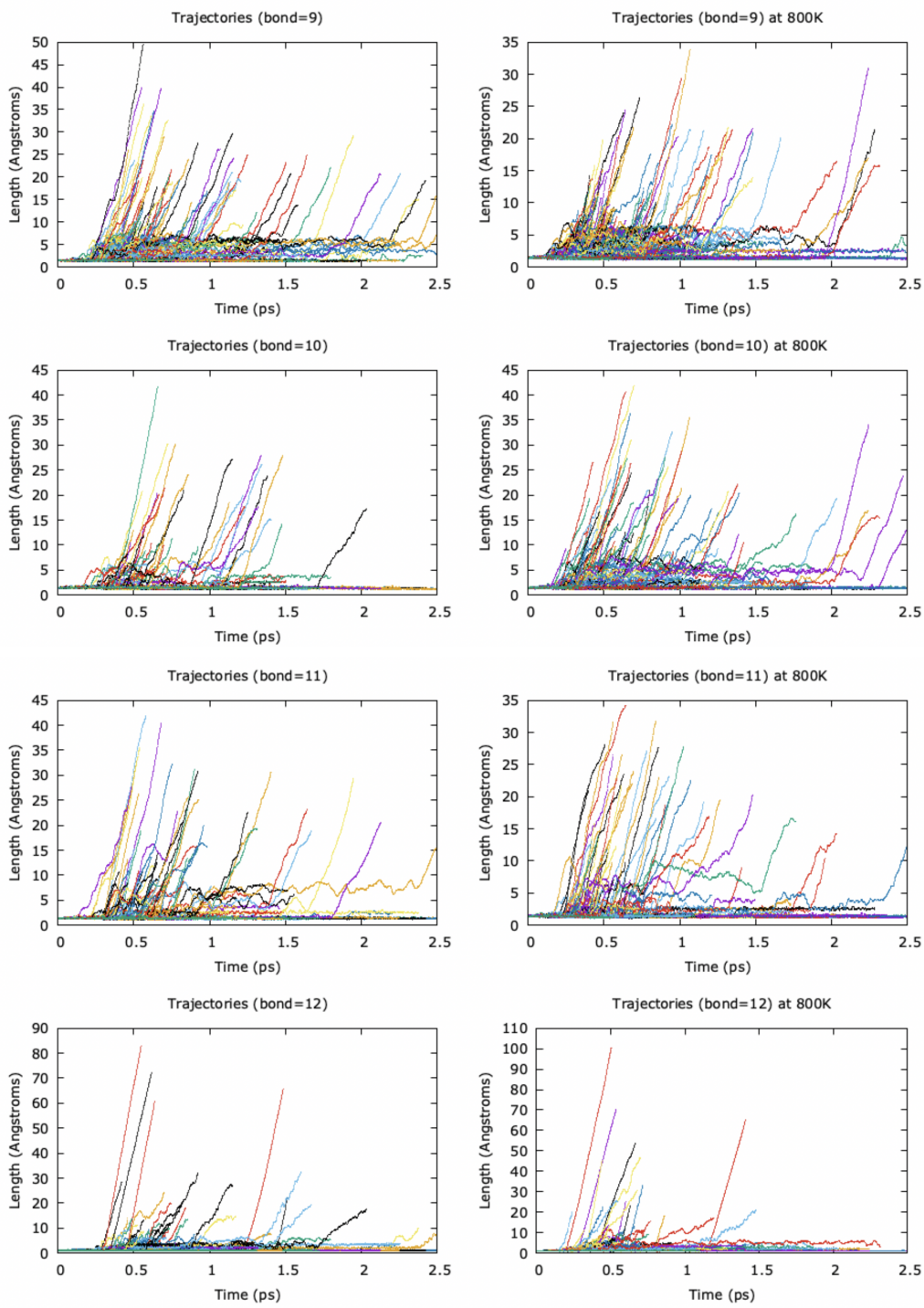


Figure S5: Bond trajectories (9-12) as a function of time after instantaneous ionization (up to 3 levels of cascaded fragmentation) of naphthalene at 500 K (left column) and 800 K (right column) vaporized substrate temperature. Bond length excursions beyond the equilibrium values (1.37-1.42$\mathring{A}$) indicate a dissociated bond.

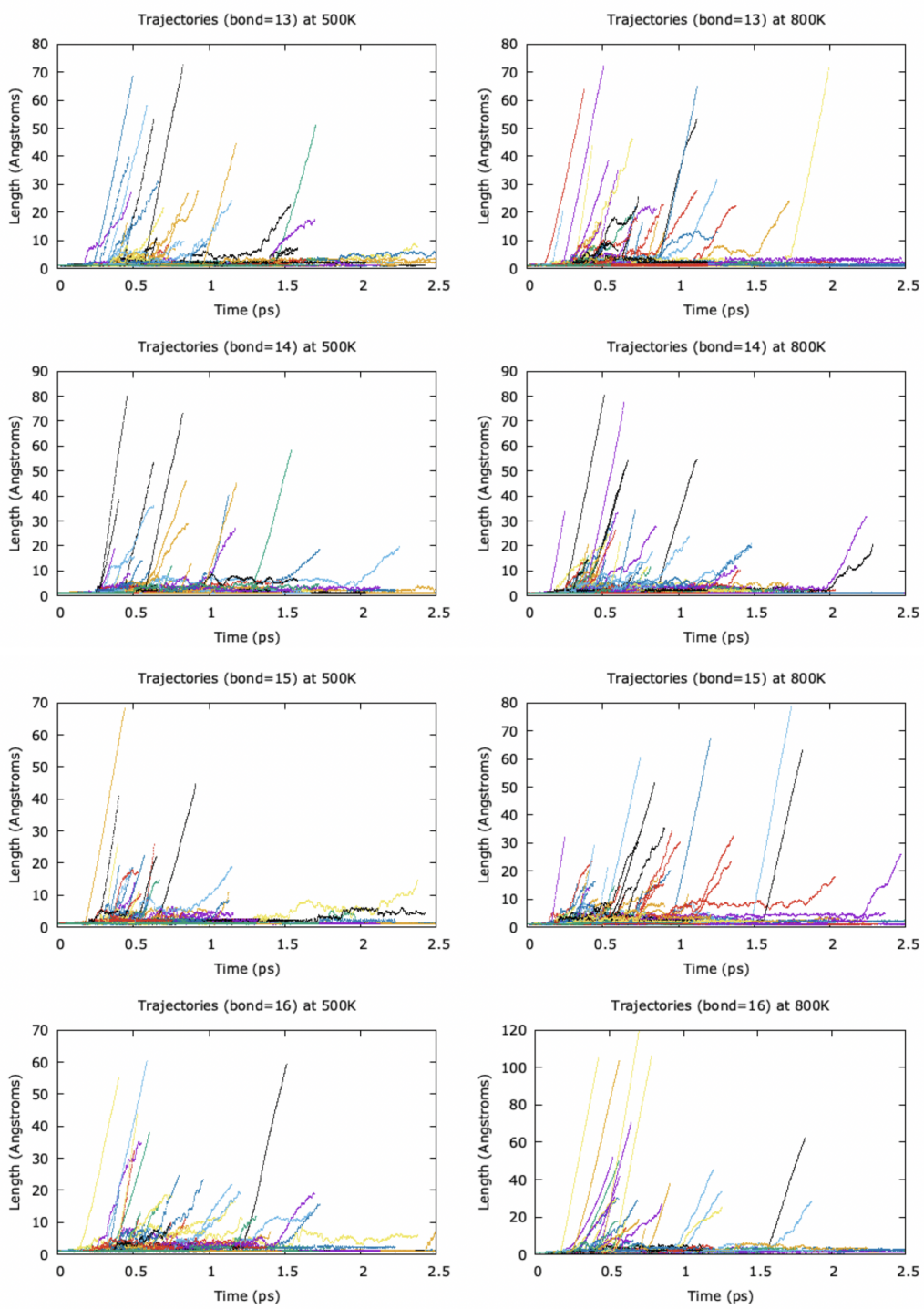


Figure S6: Bond trajectories (13-16) as a function of time after instantaneous ionization (up to 3 levels of cascaded fragmentation) of naphthalene at 500 K (left column) and 800 K (right column) vaporized substrate temperature. Bond length excursions beyond the equilibrium values (1.37-1.42$\mathring{A}$) indicate a dissociated bond.

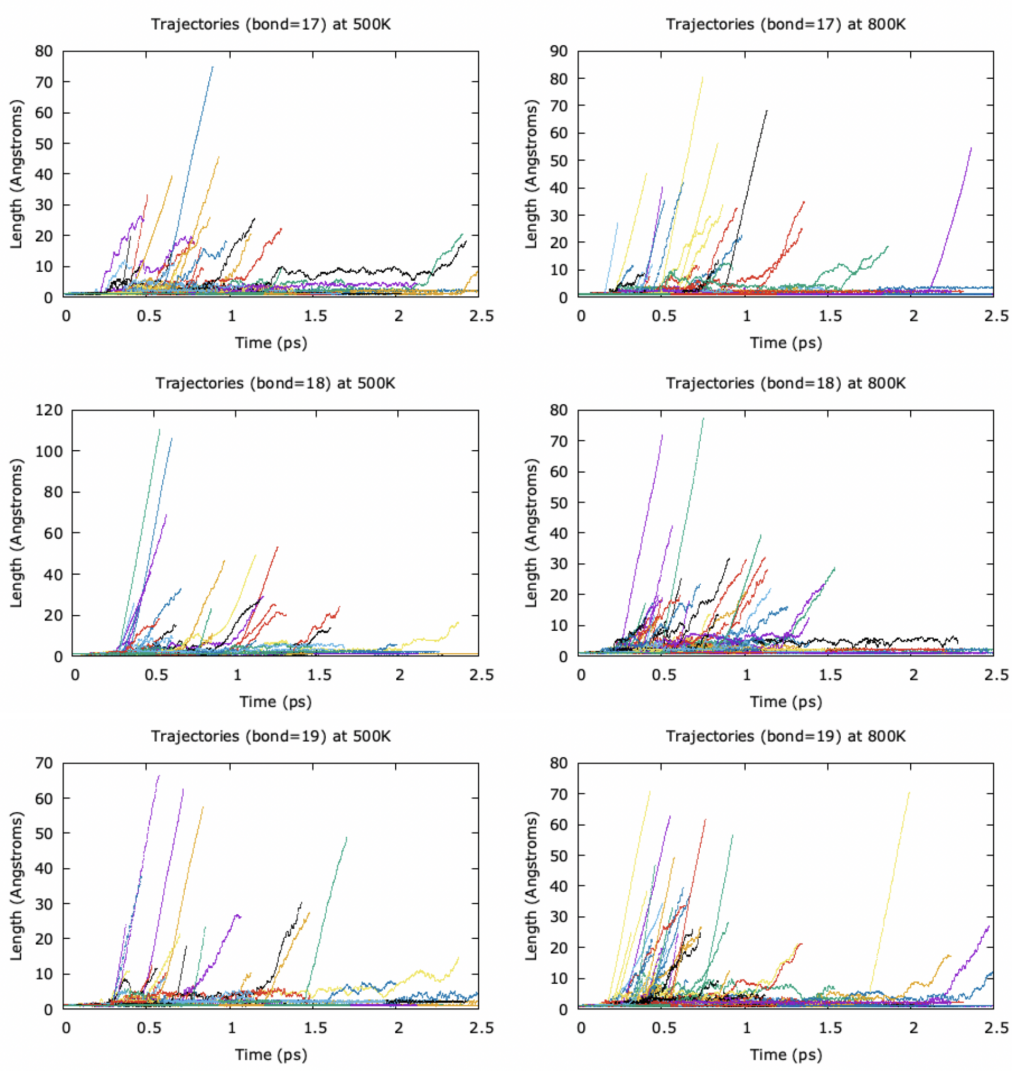


Figure S7: Bond trajectories (17-19) as a function of time after instantaneous ionization (up to 3 levels of cascaded fragmentation) of naphthalene at 500 K (left column) and 800 K (right column) vaporized substrate temperature. Bond length excursions beyond the equilibrium values (1.37-1.42$\mathring{A}$) indicate a dissociated bond.

The results shown in Figure **??**, left and right, of the manuscript reached between 215 and 224 theoretical counts in 100% signal, an experimental/theoretical MAD between 3.23925773768669 and 1.82124223544914, and a composite matching score of 0.738 and 0.768, respectively. The corresponding bond ionization trajectories are shown in Figures S2-S7 show the lower vulnerability to cascaded fragmentation from C-H bonds and increased vulnerability to cascaded fragmentation from bonds in benzene rings perimeter.

ReaxFF reactive force field and molecular dynamics

In order to describe the bond-dissociation process during HVI of molecules on walls, we use the bond-order based ReaxFF reactive molecular dynamics method, summarized graphically in Figure S8. ReaxFF provides nearly the accuracy of ground state quantum mechanics (QM) for describing reactive processes at the expense of conventional force fields. It enables the simulation of reactivity, diffusion, material decohesion and fragmentation, and phase transitions, which are essential to capturing the gas or surface chemistry and transport of molecular species during HVI events across the spacecraft and instrument surfaces. ReaxFF is used here in RMD to explore the relationship between velocity of impact and energetic fractionation in HV collisions between the different gas phase neutral molecules of interest. In ReaxFF, charges are allowed to change as bonds are formed or broken and they operate between all atoms, not just non-bonded ones, van der Waals (vdW) interactions are included between all atoms (not just non-bonded atoms), allowing the valence bonding interaction to be monotonically attractive since the vdW inner wall balances the bond attraction, all valence interactions depend on the bond order and go to zero as the bonds are broken, and all parameters are obtained directly and systematically from QM.

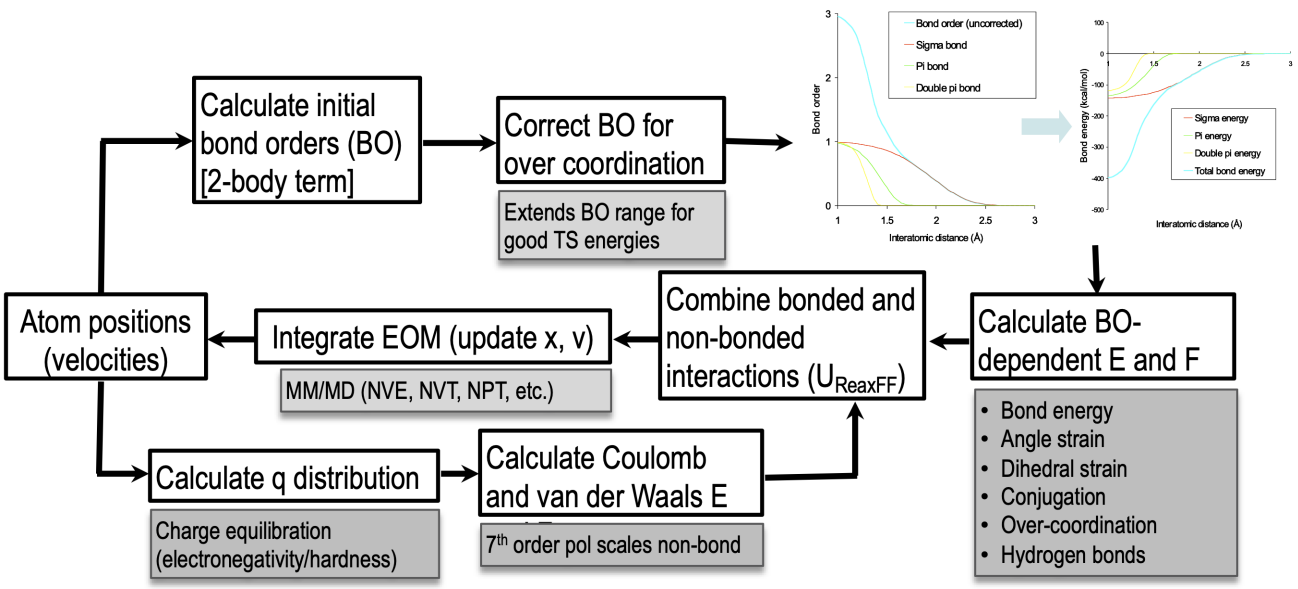


Figure S8: Classical harmonic-based potentials fail to properly capture bond breaking/formation. ReaxFF allows smooth potential energy surface transitions during chemical reactions (through bond-order corrections). Bond-order is directly associated to interatomic distances and energies for single, double and triple bonds.

ReaxFF bond curves for the different pairwise interactions in amino and fatty acids

Figure S9 compares QM predicted bond curves at different bond lengths using CI, experimental energies, and ReaxFF-predicted values for the covalent interactions involved in amino and fatty acids, longer water-insoluble fatty acids and mixed aromatic and aliphatic structures in this work. The reaction coordinate scans listed in Table 1 of the manuscript were performed at intervals of 0.05 Åusing the LAMMPS[4] molecular dynamics code.

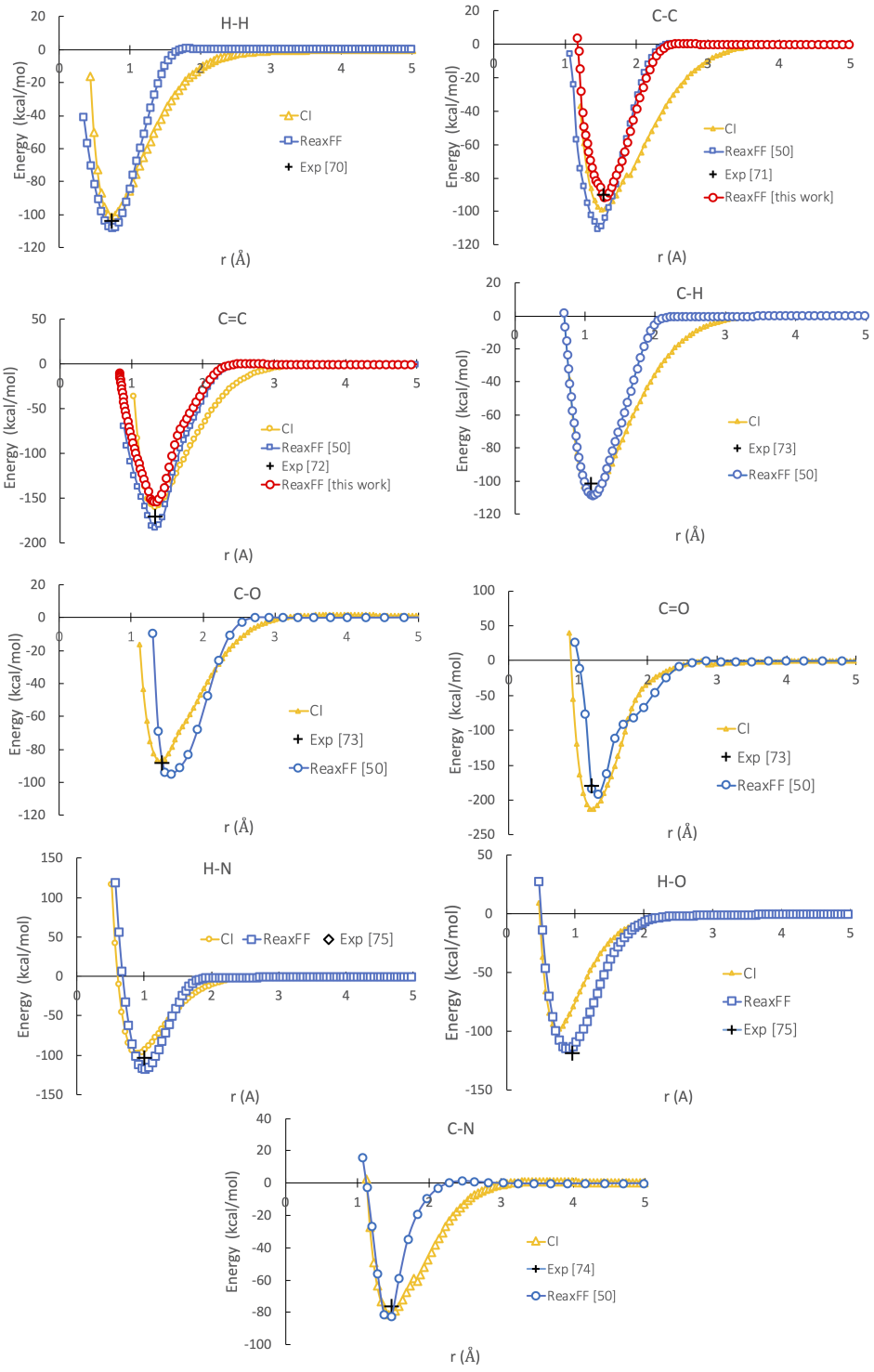


Figure S9: Bond curves resulting from the ReaxFF force field used in this work accurately reproduce equilibrium bond lengths and near-equilibrium potential energy curvatures, dissociation barriers, and inner-walls from QM at the level of configuration interaction (CI), as well as experimental bond dissociation energies and bond lengths. Blue curves correspond to the results obtained with the original ReaxFF force field in,[6] while the red curves show the specific cases affected by the change of $D_E^{\sigma} = 128.2$ for C-C bonds, i.e. C-C and C=C. Bond scans were performed on the corresponding reaction coordinates listed in Table 1 of the manuscript.

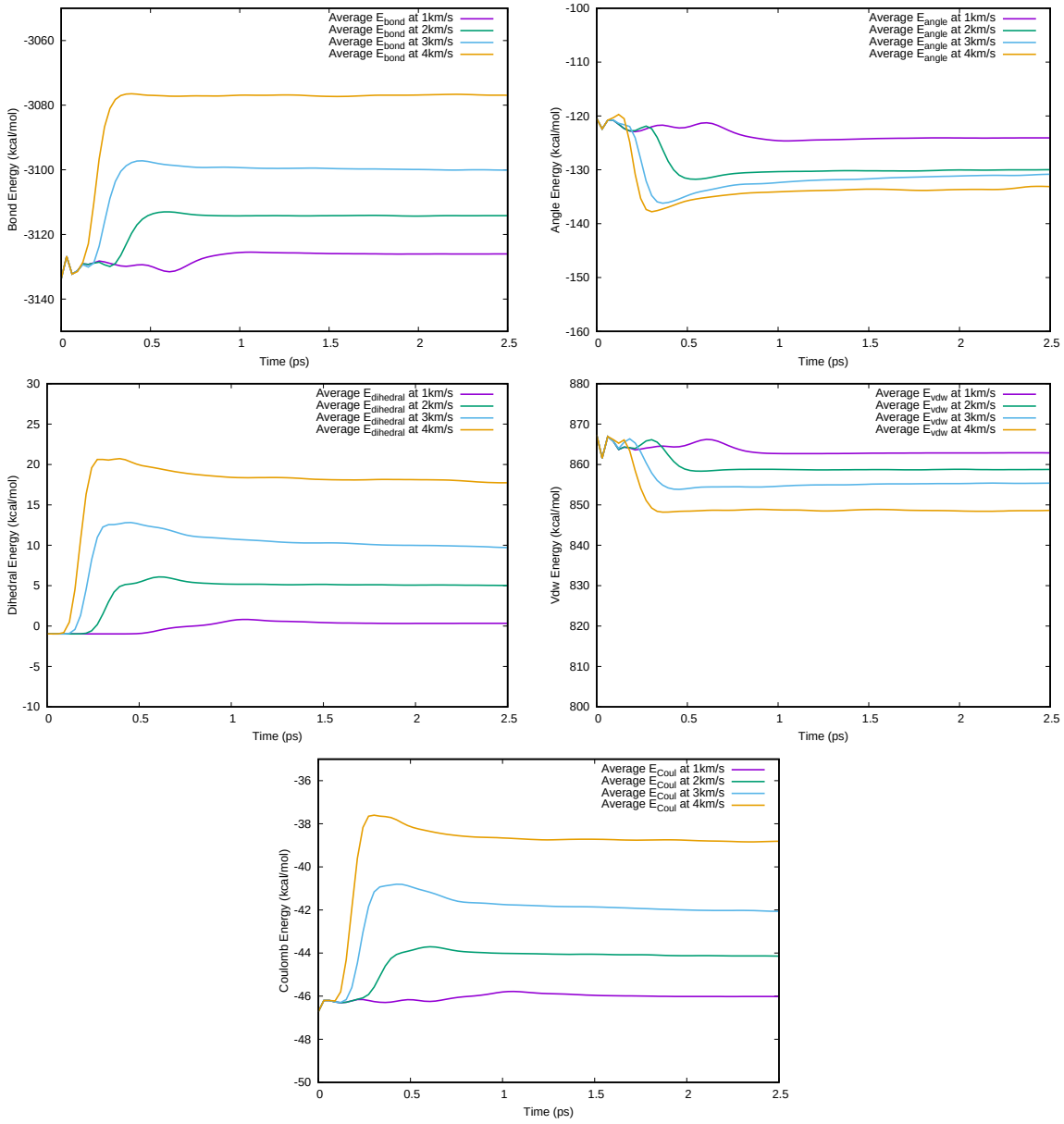


Figure S10: Covalent and non-bonded energies during HVI of molecules (naphthalene at $v = 1 - 4km/s$ depicted; impact occurs ≈0.1 ps).

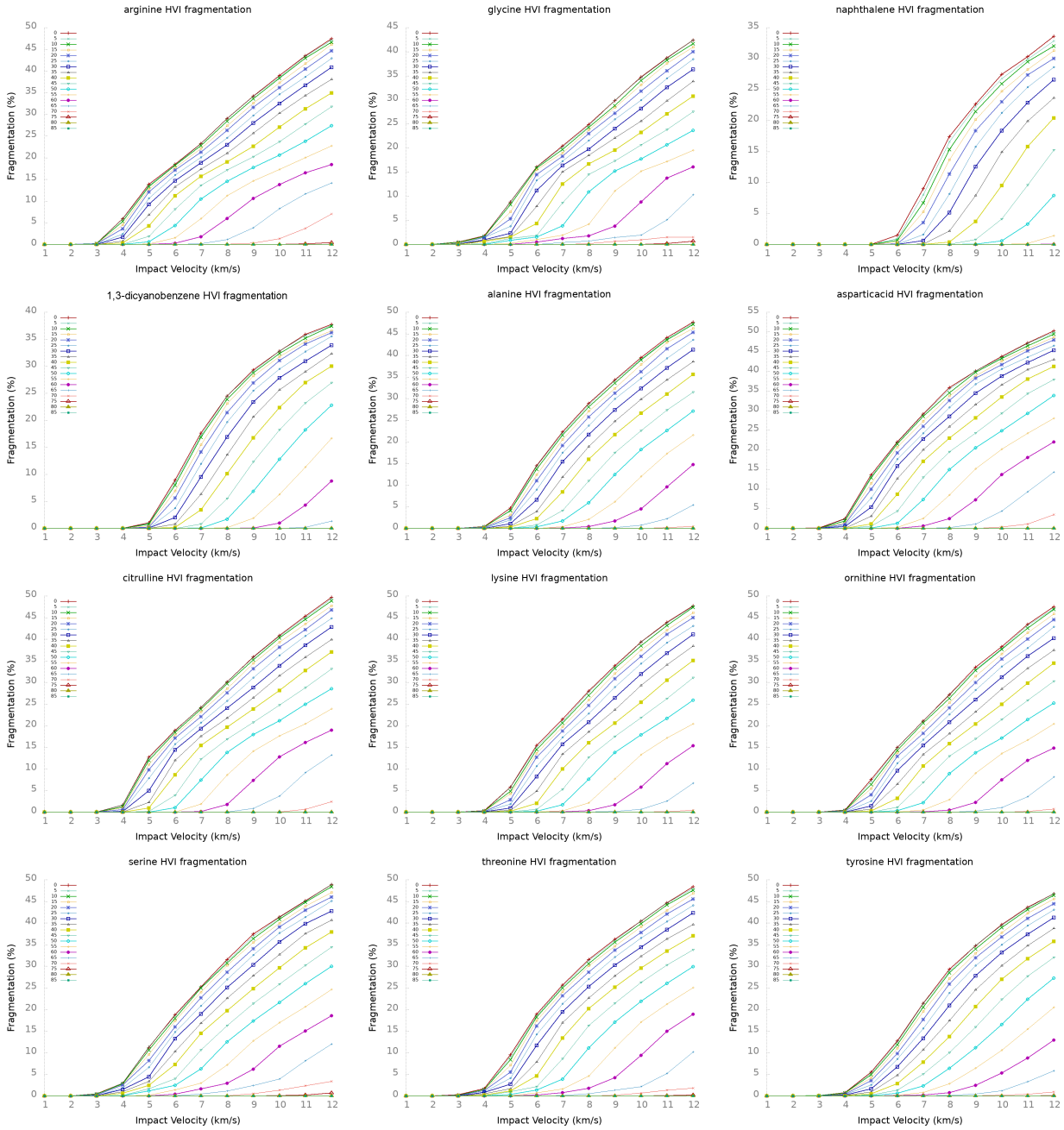


Figure S11: Molecular fragmentation percentage as a function of velocity and impact angles (represented by each line) for the different species listed in Table S1 and S3.

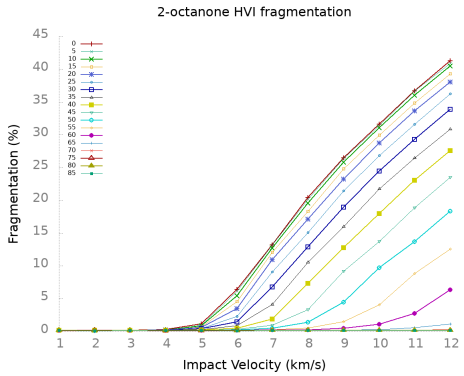
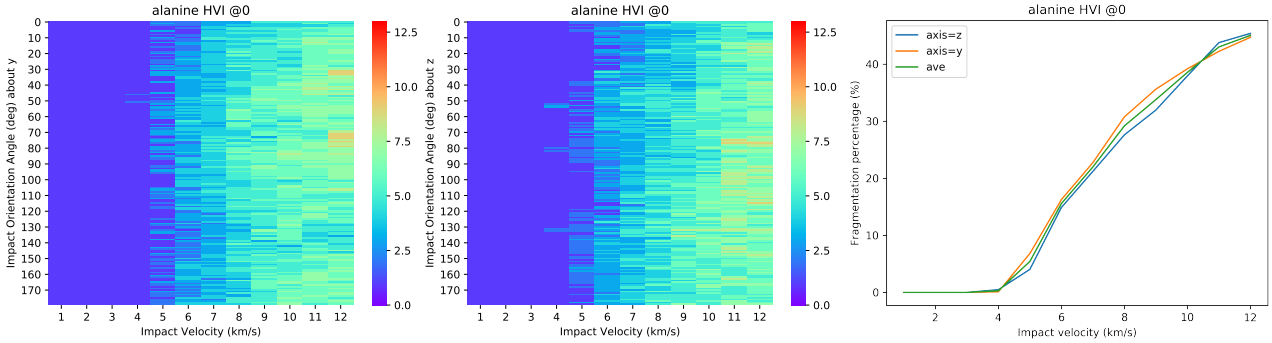
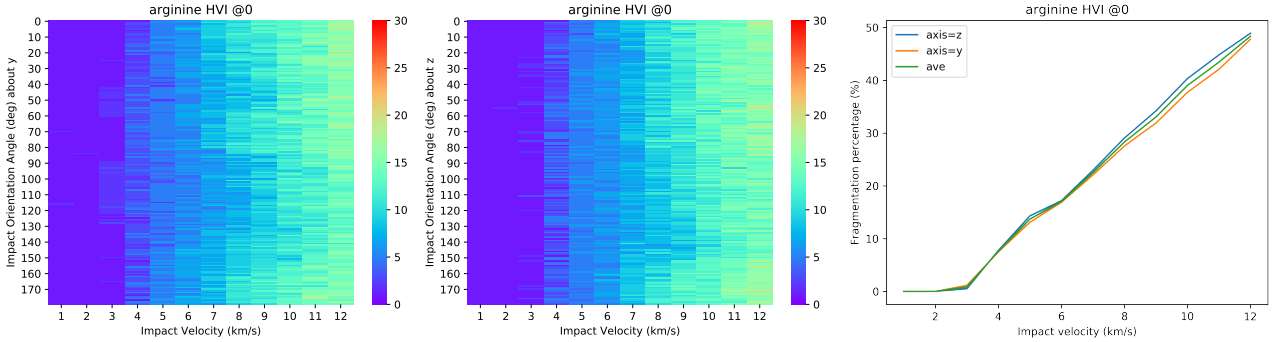
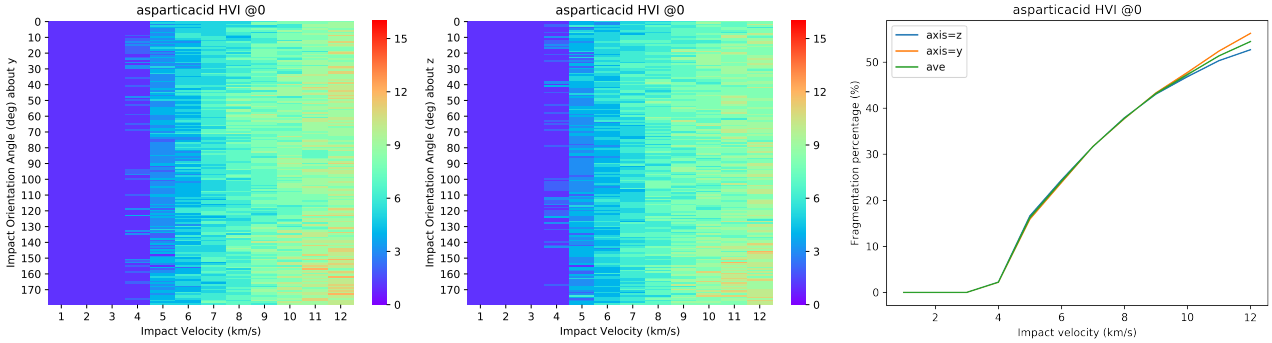
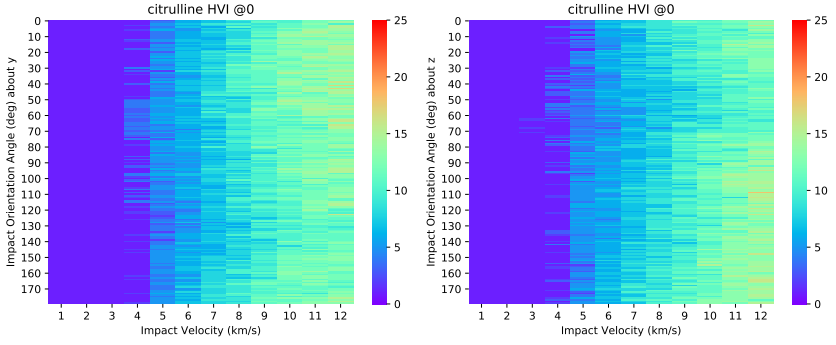


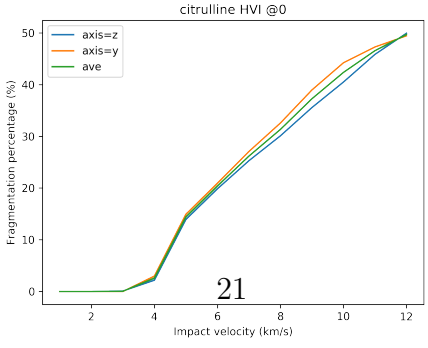
Figure S12: Molecular fragmentation percentage as a function of velocity and impact angles (represented by each line) for 2-octanone.

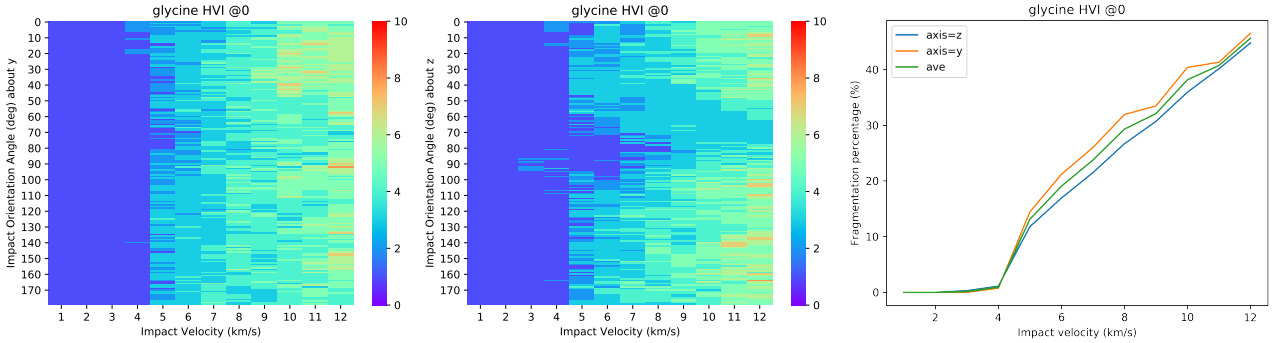

alanine HVI @0
Impact Orientation Angle (deg) about y
Impact Velocity (km/s)
alanine HVI @0
Impact Orientation Angle (deg) about z
Impact Velocity (km/s)
alanine HVI @0
axis=z
axis=y
ave
Fragmentation percentage (%)
Impact velocity (km/s)


arginine HVI @0
Impact Orientation Angle (deg) about y
Impact Velocity (km/s)
arginine HVI @0
Impact Orientation Angle (deg) about z
Impact Velocity (km/s)
arginine HVI @0
axis=z
axis=y
ave
Fragmentation percentage (%)
Impact velocity (km/s)


asparticacid HVI @0
Impact Orientation Angle (deg) about y
Impact Velocity (km/s)
asparticacid HVI @0
Impact Orientation Angle (deg) about z
Impact Velocity (km/s)
asparticacid HVI @0
axis=z
axis=y
ave
Fragmentation percentage (%)
Impact velocity (km/s)

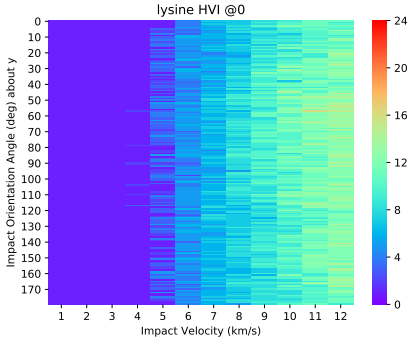

citrulline HVI @0
Impact Orientation Angle (deg) about y
Impact Velocity (km/s)
citrulline HVI @0
Impact Orientation Angle (deg) about z
Impact Velocity (km/s)

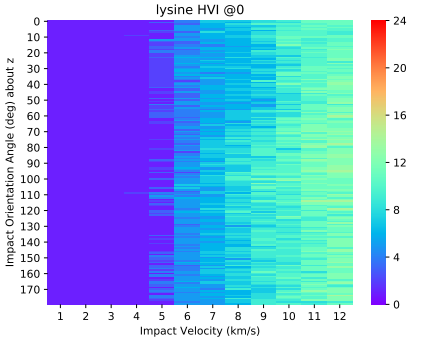

citrulline HVI @0
axis=z
axis=y
ave
Fragmentation percentage (%)
Impact velocity (km/s)

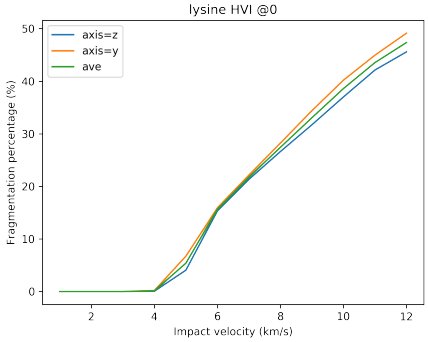

glycine HVI @0
Impact Orientation Angle (deg) about y
Impact Velocity (km/s)

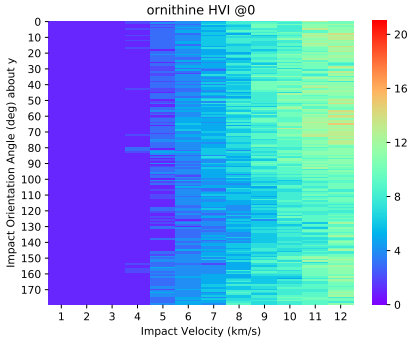
glycine HVI @0
Impact Orientation Angle (deg) about z
Impact Velocity (km/s)

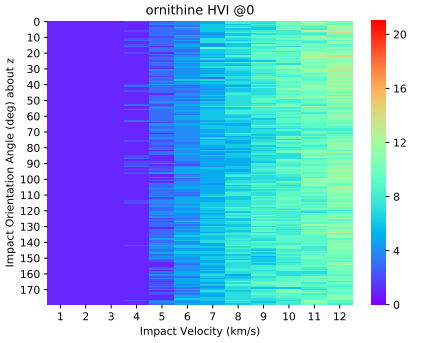
glycine HVI @0
axis=z
axis=y
ave
Fragmentation percentage (%)
Impact velocity (km/s)

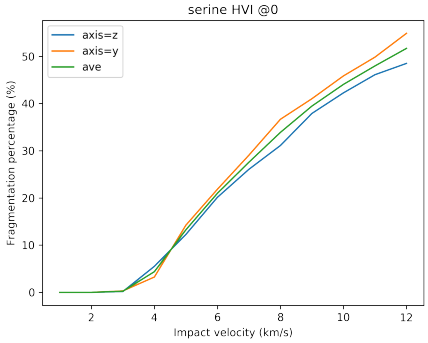

lysine HVI @0
Impact Orientation Angle (deg) about y
Impact Velocity (km/s)

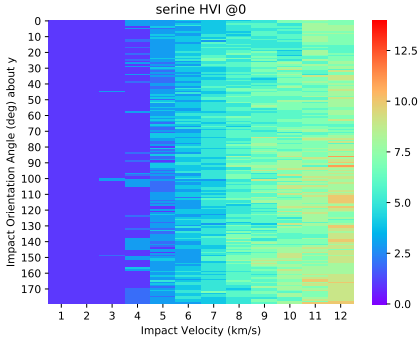

lysine HVI @0
Impact Orientation Angle (deg) about z
Impact Velocity (km/s)

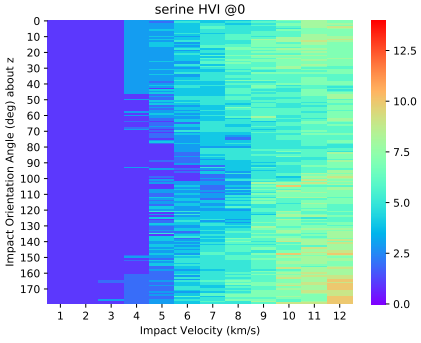

lysine HVI @0
axis=z
axis=y
ave
Fragmentation percentage (%)
Impact velocity (km/s)

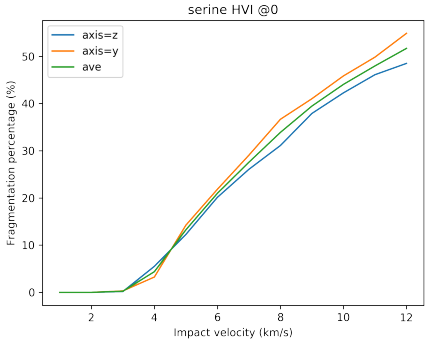

ornithine HVI @0
Impact Orientation Angle (deg) about y
Impact Velocity (km/s)


ornithine HVI @0
Impact Orientation Angle (deg) about z
Impact Velocity (km/s)


serine HVI @0
axis=z
axis=y
ave
Fragmentation percentage (%)
Impact velocity (km/s)


serine HVI @0
Impact Orientation Angle (deg) about y
Impact Velocity (km/s)


serine HVI @0
Impact Orientation Angle (deg) about z
Impact Velocity (km/s)


serine HVI @0
axis=z
axis=y
ave
Fragmentation percentage (%)
Impact velocity (km/s)

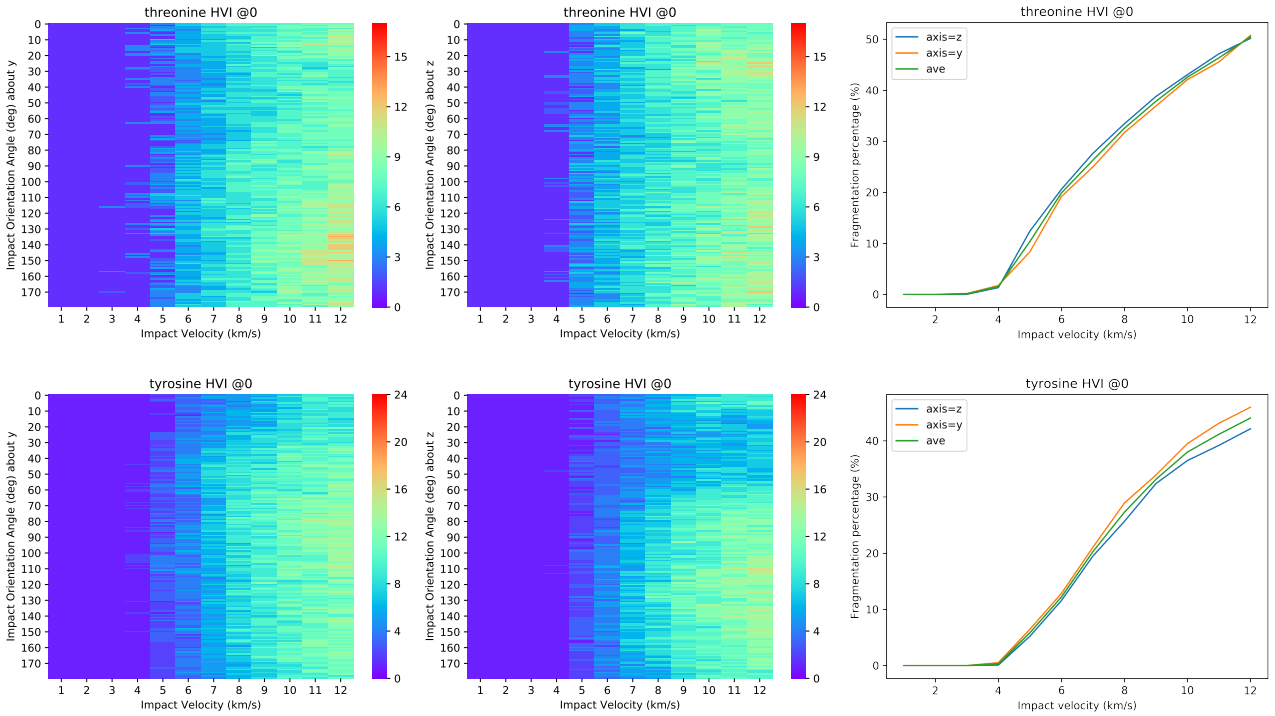


Figure S13: HVI fragmentation heatmaps at normal angle over $v = 1 - 12$ km/s for left column) changes in landing orientation about y and center column) changes in landing orientation about z. Right column xy depict the percent fragmentation per species over $v = 1 - 12$ km/s, averaged over 36 angles about y and 36 angles about z.

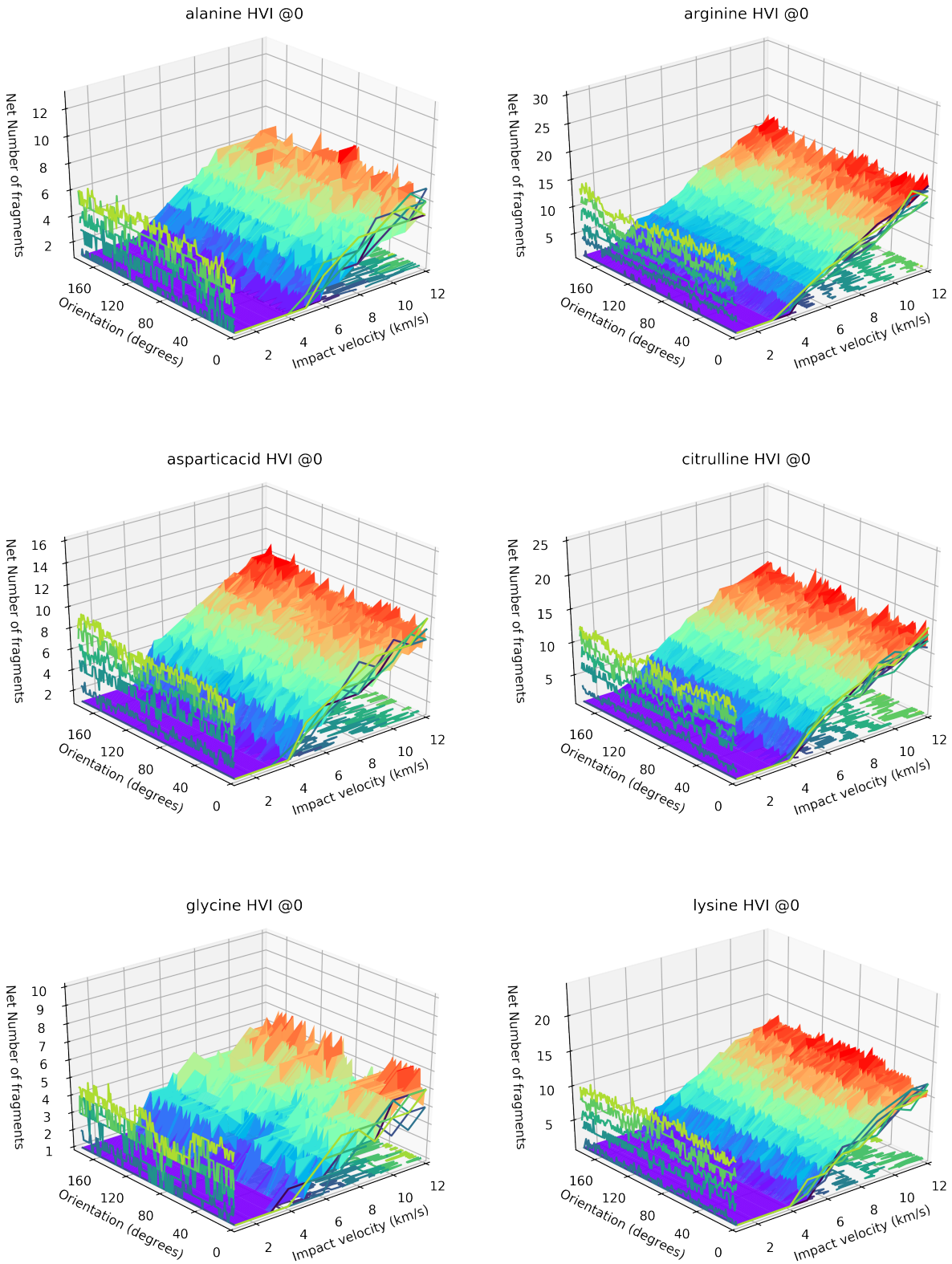


Figure S14: HVI 3D surface plot of the number of fragmentation species at normal angle over $v = 1 - 12$ km/s vs averaged changes in landing orientation..

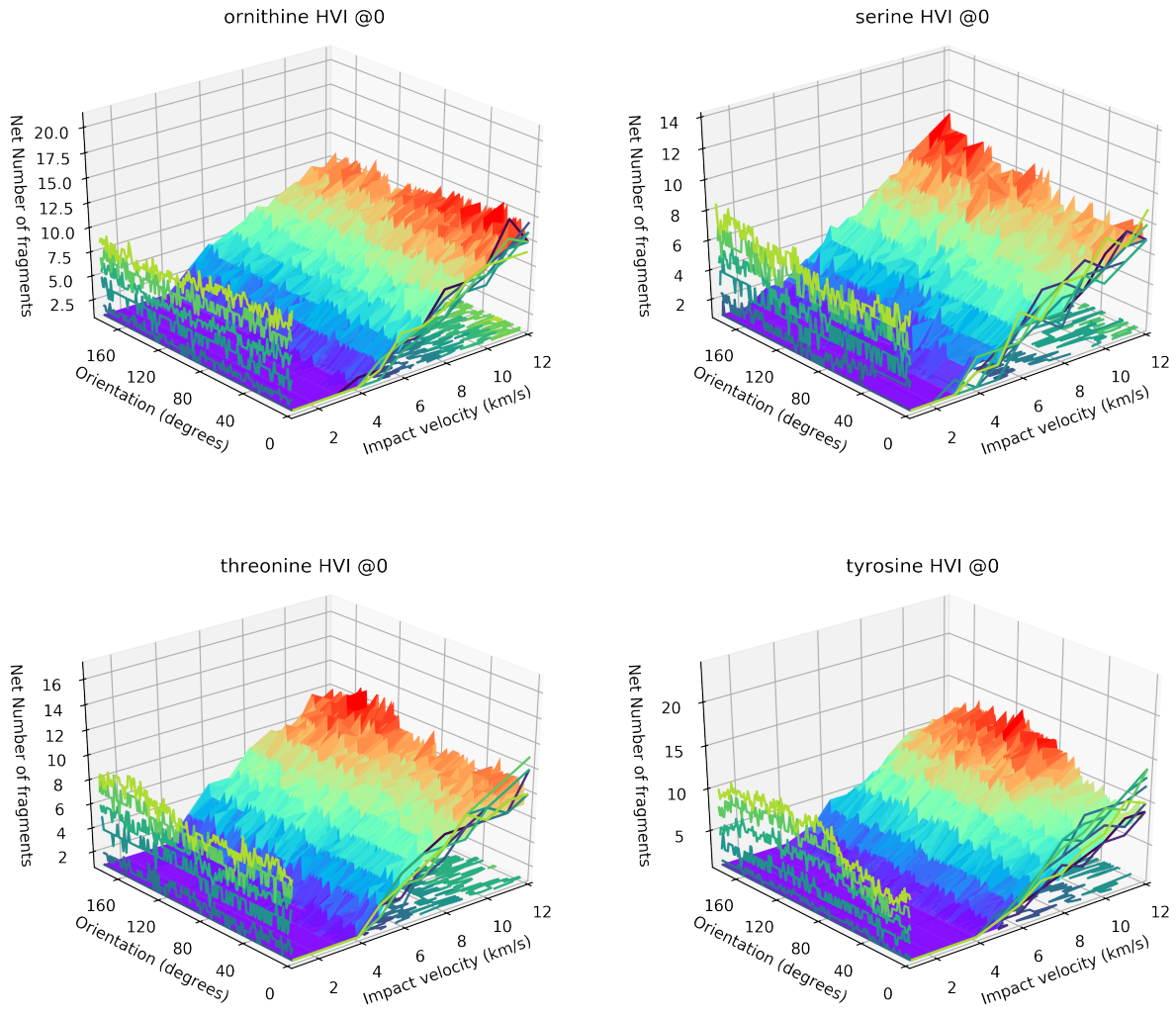


Figure S15: HVI 3D surface plot of the number of fragmentation species at normal angle over $v = 1 - 12$ km/s vs averaged changes in landing orientation.

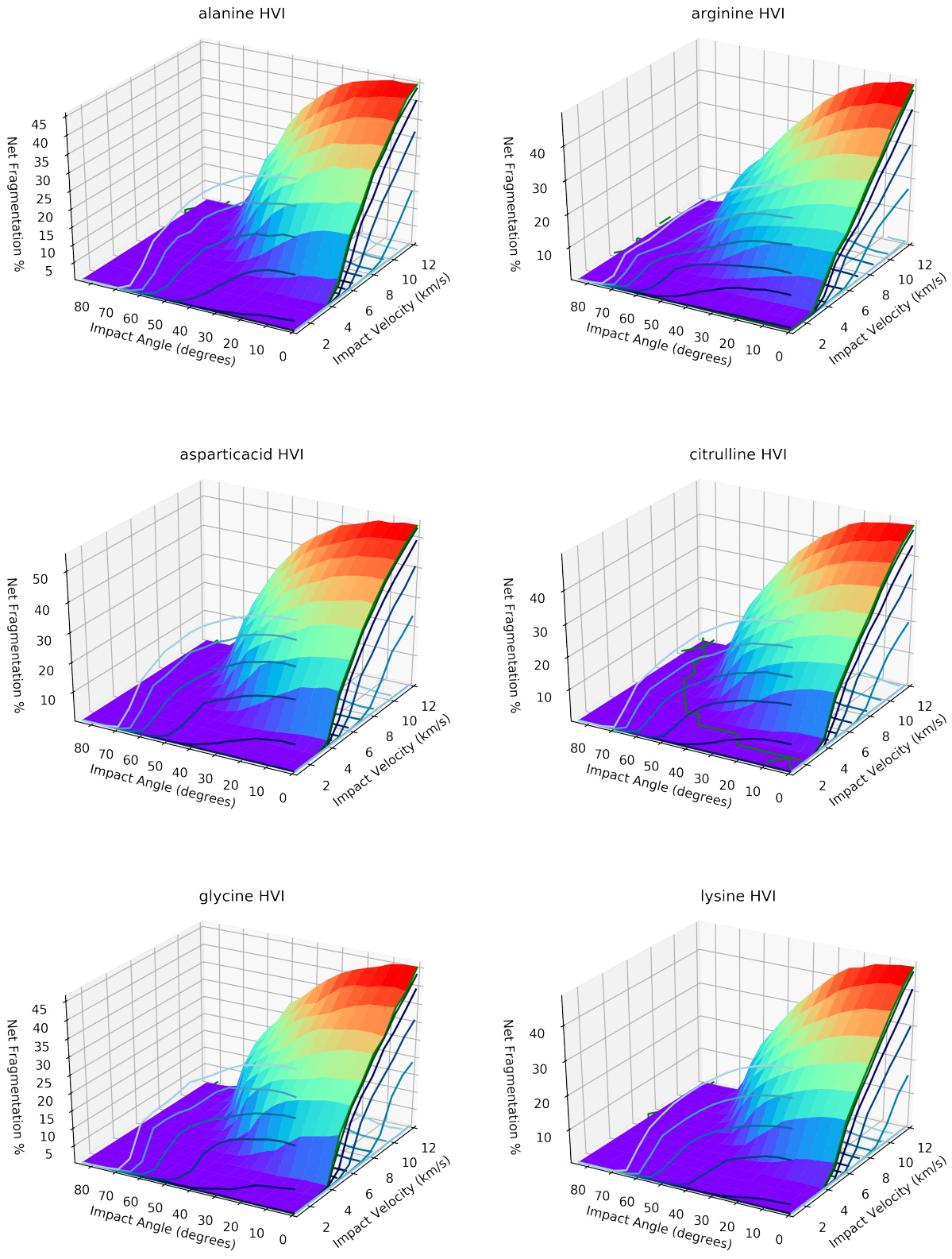


Figure S16: HVI 3D surface plot of the fragmentation percent vs impact velocity and impact angle for Ala, Arg, Asp, Cit, Gly and Lys. Every impact angle is averaged over the entire set of landing orientations, per impact velocity.

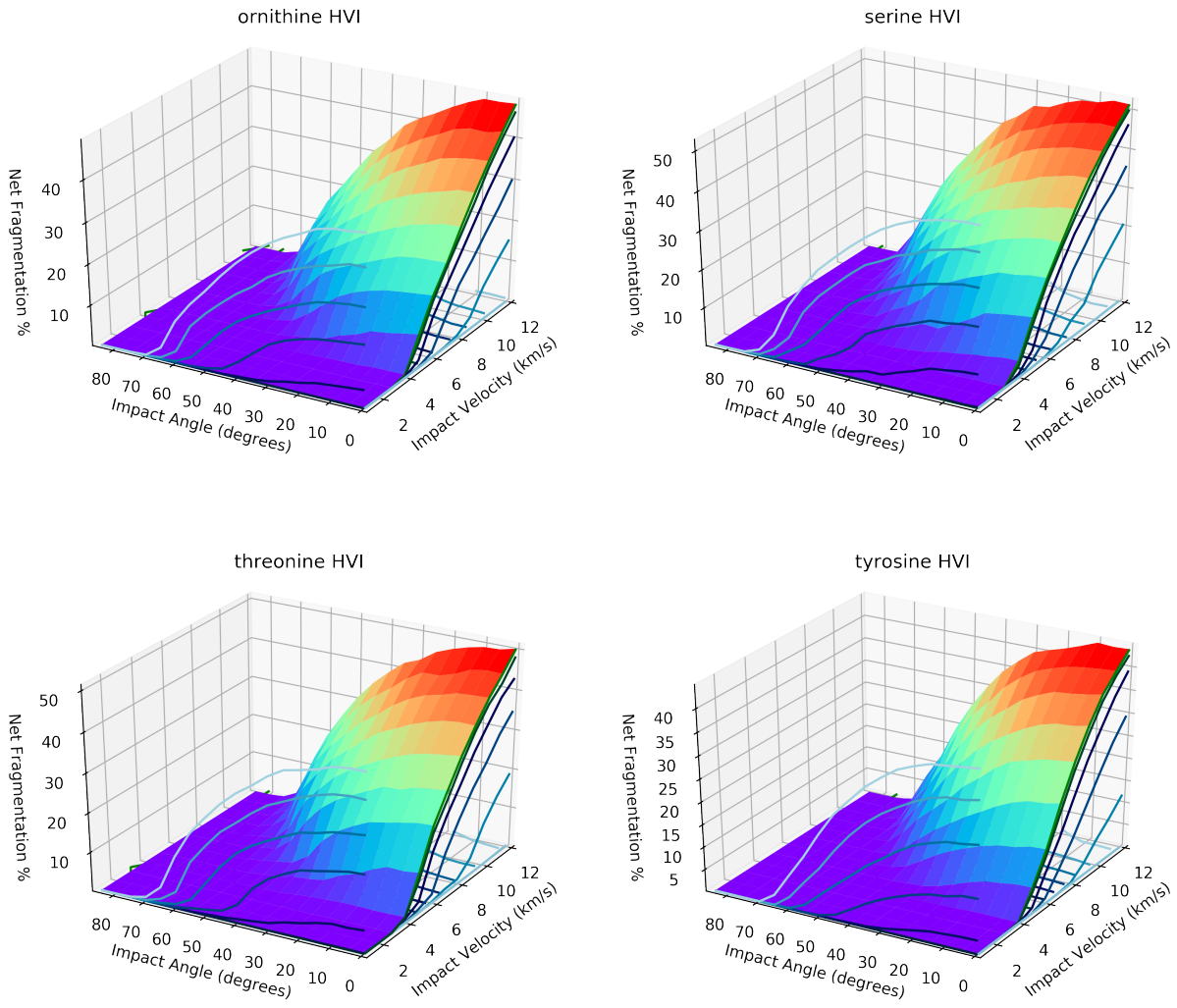


Figure S17: HVI 3D surface plot of the fragmentation percent vs impact velocity and impact angle for Orn, Ser, Thr, and Tyr. Every impact angle is averaged over the entire set of landing orientations, per impact velocity.

The triclinic unit cell used to construct the ice nanoparticles, from supercells of the former, is included below in PDB (.pdb) format.

COMPND Ice-Ih

AUTHOR GENERATED BY AJB, from J. Chem. Phys. 1, 515 (1933) J.D. Bernal and R. H. Fowler

CRYST1 7.820 7.820 7.360 90.00 90.00 120.00 P 63 c m 1

HETATM 1 O HOH 1 2.606 0.000 0.460 1.00 0.00 O2-

HETATM 2 O HOH 2 5.214 0.000 6.900 1.00 0.00 O2-

HETATM 3 H HOH 0 2.606 0.000 1.281 1.00 0.00 H1+

HETATM 4 H HOH 0 3.425 0.000 0.191 1.00 0.00 H1+

HETATM 5 H HOH 0 5.626 0.711 7.176 1.00 0.00 H1+

CONECT 1 4 3

CONECT 2 5

CONECT 3 1

CONECT 4 1

CONECT 5 2

MASTER 0 0 0 0 0 0 0 0 5 0 5 0

END

References

(1) Lide, D. R., Ed. *CRC Handbook of Chemistry and Physics*, 1st ed.; CRC Press/Taylor & Francis Group: Boca Raton, FL., 2008; Vol. 130.

(2) Bochevarov, A. D.; Harder, E.; Hughes, T. F.; Greenwood, J. R.; Braden, D. A.; Philipp, D. M.; Rinaldo, D.; Halls, M. D.; Zhang, J.; Friesner, R. A. *International Journal of Quantum Chemistry* **2013**, *113*, 2110–2142.

(3) M.S.Gordon,; M.W.Schmidt, *Advances in electronic structure theory: GAMESS a decade later*, theory and applications of computational chemistry: the first forty years ed.; Elsevier, Amsterdam, 2005; pp 1167–1189.

(4) Plimpton, S. *Journal of Computational Physics* **1995**, *117*, 1 – 19.

(5) Aktulga, H. M.; Fogarty, J. C.; Pandit, S. A.; Grama, A. Y. *Parallel Comput.* **2012**, *38*, 245259.

(6) Verlackt, C. C. W.; Neyts, E. C.; Jacob, T.; Fantauzzi, D.; Golkaram, M.; Shin, Y.-K.; van Duin, A. C. T.; Bogaerts, A. *New Journal of Physics* **2015**, *17*, 103005.

(7) Liu, L.; Liu, Y.; Zybin, S. V.; Sun, H.; Goddard, W. A. *The Journal of Physical Chemistry A* **2011**, *115*, 11016–11022, PMID: 21888351.

(8) Shinoda, W.; Shiga, M.; Mikami, M. *Phys. Rev. B* **2004**, *69*, 134103.

(9) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Phys. Chem.* **1994**, *101*, 4177–4189.

(10) Parrinello, M.; Rahman, A. *Journal of Applied Physics* **1981**, *52*, 7182–7190.

(11) Grimme, S. *Angewandte Chemie International Edition* **2013**, *52*, 6306–6312.